




Predictors of mortality of COVID-19 in the general population and nursing homes

Pedro Pablo España^{1,6} · Amaia Bilbao^{2,4,5} · Susana García-Gutiérrez^{3,4,5} · Iratxe Lafuente^{3,4,5} · Ane Anton-Ladislao^{3,4,5} · Ane Villanueva^{3,4,5} · Ane Uranga^{1,6} · Maria Jose Legarreta^{3,4,6} · Urko Aguirre^{3,4,5} · Jose Maria Quintana^{3,4,5}  · the COVID-19-Osakidetza Working group

Received: 10 September 2020 / Accepted: 27 November 2020 / Published online: 5 January 2021
© Società Italiana di Medicina Interna (SIMI) 2021

Abstract

The factors that predispose an individual to a higher risk of death from COVID-19 are poorly understood. The goal of the study was to identify factors associated with risk of death among patients with COVID-19. This is a retrospective cohort study of people with laboratory-confirmed SARS-CoV-2 infection from February to May 22, 2020. Data retrieved for this study included patient sociodemographic data, baseline comorbidities, baseline treatments, other background data on care provided in hospital or primary care settings, and vital status. Main outcome was deaths until June 29, 2020. In the multivariable model based on nursing home residents, predictors of mortality were being male, older than 80 years, admitted to a hospital for COVID-19, and having cardiovascular disease, kidney disease or dementia while taking anticoagulants or lipid-lowering drugs at baseline was protective. The AUC was 0.754 for the risk score based on this model and 0.717 in the validation subsample. Predictors of death among people from the general population were being male and/or older than 60 years, having been hospitalized in the month before admission for COVID-19, being admitted to a hospital for COVID-19, having cardiovascular disease, dementia, respiratory disease, liver disease, diabetes with organ damage, or cancer while being on anticoagulants was protective. The AUC was 0.941 for this model's risk score and 0.938 in the validation subsample. Our risk scores could help physicians identify high-risk groups and establish preventive measures and better follow-up for patients at high risk of dying.

ClinicalTrials.gov Identifier: NCT04463706

Keywords COVID-19 · Cohort study · Prognostic factors · Mortality · Clinical prediction rules

Details of the COVID-19-Osakidetza Working group is given in the acknowledgement section.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11739-020-02594-8>.

✉ Jose Maria Quintana
josemaria.quintanalopez@osakidetza.eus

¹ Osakidetza/Basque Health Service, Respiratory Service, Galdakao University Hospital, Galdakao, Bizkaia, Spain

² Osakidetza/Basque Health Service, Research Unit, Basurto University Hospital, Bilbao, Bizkaia, Spain

³ Osakidetza/Basque Health Service, Research Unit, Galdakao University Hospital, Galdakao, Bizkaia, Spain

Introduction

The outbreak of disease caused by the novel severe respiratory syndrome coronavirus 2 (SARS-CoV-2) was declared a pandemic by the World Health Organization on March 11, 2020 [1]. So far, the factors that put individuals at a higher risk of poor outcomes from COVID-19 are poorly understood. Previously reported studies have weaknesses in

⁴ Health Service Research Network on Chronic Diseases (REDISSEC), Bizkaia, Spain

⁵ Kronikgune Institute for Health Services Research, Barakaldo, Spain

⁶ BioCruces-Bizkaia Health Institute, Bizkaia, Spain

their design, [2–6] including relatively small sample sizes and analysis of single-center data, and most only consider hospitalized patients, without providing data on the general population [7], or more frail groups, in particular, those institutionalized in nursing homes. Furthermore, these studies have focused on disease severity, rather than mortality.

Although it seems clear that older patients with chronic medical conditions are at a higher risk of being infected [8], it is necessary to verify the clinical characteristics and case fatality of COVID-19 in countries with different demographic characteristics to generalize these findings. Furthermore, some drugs received as concomitant treatments for chronic medical conditions at baseline may be affecting the course of the disease. Early identification of people at the highest risk of deteriorating would help physicians detect vulnerable groups, to administer preventive therapies or vaccinations and minimize further spread of the infection. Effective patient risk stratification is essential to optimize care and the use of healthcare resources. We need "prediction models" for the general population, to detect target groups to guide medical staff in triaging patients for allocating finite healthcare resources [9].

The aim of this study was to identify factors associated with risk of death among patients diagnosed with COVID-19 in the Basque Country and thereby create and validate prediction scores.

Methods

This is a retrospective study of a cohort of patients diagnosed with COVID-19 in the Basque Country based on data from the electronic database and health records of the Basque Health Service-Osakidetza. This Spanish region has a population of 2,178 million people, the vast majority of whom are entitled to healthcare under Osakidetza. The Basque Health System is divided into 13 integrated healthcare organizations (IHOs), gathering all primary and hospital care resources in given areas under the same administrative management.

All patients included in this study were residents in the Basque Country with SARS-CoV-2 infection that was laboratory-confirmed by a positive result on the reverse transcriptase–polymerase chain reaction assay for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or a positive IgM or IgG antibody test performed due to symptoms suggestive of the disease or having had contact with a positive case, from February to May 22, 2020. No patients were excluded. The study protocol was approved by the Ethics Committee of the Basque Country (reference PI2020059). All patient data were kept confidential.

All data on patients under the care of Osakidetza are recorded in a unified electronic database. Analysts retrieved data from all positive cases detected during the study

period including sociodemographic data (age, sex, place of residence), baseline comorbidities (all those considered in Charlson's Comorbidity index [10], plus angina, arrhythmia, arterial hypertension, dyslipidemia, asthma, bronchiectasis, cystic fibrosis, interstitial lung disease, lymphoma, leukemia, coagulopathy, inflammatory bowel disease, gastrointestinal bleeding), baseline treatments (based on the Anatomical, Therapeutic, Chemical [ATC] classification system) [11–13], other background data concerning care provided in hospital or primary care settings including dates of hospital admission and discharge, and whether patients were admitted to an intensive care unit (ICU), and vital status. Comorbidities were identified based on the International Statistical Classification of Diseases and Related Health Problems (ICD) ICD-9 or 10 codes in patients' records at baseline [14].

We grouped comorbidities in the following way: cardiovascular diseases (including myocardial infarction, angina, arrhythmia, congestive heart failure, and peripheral vascular disease); cerebrovascular disease, hemiplegia and/or paraplegia; arterial hypertension; dyslipidemia; dementia; respiratory disease (chronic obstructive pulmonary disease [COPD], bronchiectasis, chronic bronchial infection); asthma; liver disease (mild liver, moderate or severe liver disease); diabetes (diabetes with/without organ damage), kidney disease, cancer (malignant tumor, metastatic solid tumor, lymphoma); rheumatic disease; peptic ulcer; inflammatory bowel disease; and coagulopathies.

Regarding baseline medication, we selected drugs based on ATC codes [11]. Baseline treatment was defined as any drugs prescribed before the patient was diagnosed with SARS-CoV-2 infection and had no end date. Online Resource Table 1 provides summary information about the grouping of baseline treatments included in our study.

Data identifying people living in a nursing home were obtained from the Basque Health Department. The only outcome of the study was death during the study period. All patients were followed up until June 29, 2020.

Additionally, we studied the validity of our electronic database by comparing the data obtained from it for subset variables with the information provided by a group of trained reviewers who retrieved the same information, item by item, from the electronic health records of the same patients. The results of this sub-study are summarized in Online Resource Table 2. The rate of agreement was at least 93.33%, and in most cases, there was full agreement (100%).

Statistical analysis

As described in the literature on community-acquired pneumonia, we found differences in baseline characteristics and clinical course between patients from nursing homes and those from the general population. For this reason, we

divided the sample (general population or nursing home residents) and performed all the analyses separately for these two groups.

Additionally, each sample was randomly divided into two subsamples, for derivation and validation purposes (60% and 40% of the entire sample, respectively).

Descriptive statistics were generated including frequency tables for categorical variables and means, standard deviations (SDs), medians and interquartile ranges (IQRs) for continuous variables. Patient characteristics were compared between the two subsamples (derivation vs. validation) using Chi-square or Fisher's exact tests for categorical variables, and Student's *t* test or nonparametric Wilcoxon tests for continuous variables.

Univariable logistic regression models were first built using the derivation samples to identify the significance of each potential risk factor. In these models, mortality was used as the dependent variable and all candidate predictive variables (described previously) as the independent variables. Then, independent variables with a $p < 0.20$ in the univariable analyses were considered potential independent variables in the multivariable analysis, for which multilevel analyses with generalized estimated equations were performed considering the IHO. Potential interactions between variables were also examined. In the final models, only factors with $p < 0.05$ were retained. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated. The predictive accuracy of the model was determined by calculating areas under the ROC curve (AUCs) [15].

To develop each predictive risk score, we first assigned a weight to each risk factor in relation to each β parameter based on the multilevel models. Then, we summed the weights of each of the risk factors of a patient, higher scores indicating a greater likelihood of death. The predictive accuracy of the mortality risk score was assessed using the AUC [15], in both derivation and validation samples (external validation). Furthermore, we sought to validate the risk scores by *K*-fold cross-validation [16, 17], which uses part of the available data of the derivation sample to fit the model, and a different part to test it (internal validation) [16].

Having developed the mortality risk score, we categorized the score into different levels of risk. For each category, the optimal thresholds on the continuous risk scores were determined with the *catpredi* function of the R package *CatPredi* using the genetic algorithm [18]. The performance of each risk classification was studied by comparing the mortality rate between categories and using the multilevel analyses with generalized estimated equations and AUCs, in both derivation and validation samples. Finally, Kaplan–Meier curves were constructed for each risk category, and comparisons were performed with the log-rank test.

All effects were considered significant at $p < 0.05$. All statistical analyses were performed using SAS for Windows,

version 9.4 (SAS Institute, Carey, NC), and R© version 4.0.2.

Results

In this study, 18,768 people were classified as having COVID-19 based on test results. Overall, 5775 (30.77%) were admitted to a hospital during the study period and 448 (2.39%) were admitted to an ICU. This total included 3567 patients from nursing homes; their mean age was 84.15 (SD: 10.88) years and 69.67% were women. In this sample, 871 (24.42%) were admitted to a hospital during the study period and 15 (0.42%) to an ICU. In the general population sample ($n = 15,201$), the mean age was 53.77 (SD: 17.50), 59.55% were women; 4904 (32.26%) were admitted to a hospital during the study period and 433 (2.85%) to an ICU. Among the general population, 10,002 patients (65.80%) were positive at the pharyngeal swab test and 5199 patients (34.20%) at the IgG/IgM test; however, among nursing home residents, 71.32% were positive at the pharyngeal swab and 28.68% at the IgG/IgM test, being these differences statistically significant ($p < 0.0001$). As expected, patients in nursing homes presented a higher proportion of all kind of comorbidities, but mainly cardio and cerebrovascular, diabetes, COPD and dementia. Descriptive data of all patients are summarized in Online Resource Table 3.

The univariable analysis showed that for nursing home residents being older or male, having cardiovascular, cerebrovascular or kidney diseases, arterial hypertension or dementia, and being admitted to a hospital for COVID-19, as well as taking diuretics, cardiovascular medications or azithromycin at baseline were related to death during follow-up while taking NSAIDs, anticoagulants or lipid-lowering drugs were protective. In the case of patients from the general population, predictors of death identified in the univariable analysis were again older age, male sex, and being admitted to a hospital for COVID-19, and also having being admitted to a hospital in the previous month for any reason, most comorbidities, except asthma and inflammatory bowel disease, and all baseline treatments studied increased the risk, except the use of NSAIDs that was again protective (Table 1).

In the multivariable model of nursing home residents, predictors of mortality were being male, being older than 80 years, being admitted to a hospital for COVID-19, having cardiovascular disease, kidney disease or dementia at baseline, and the patient's IHO. On the other hand, taking anticoagulants or lipid-lowering drugs at baseline was found to be protective. The AUC was 0.754 for this model overall and 0.742 not including the IHO in the model (Table 2).

For patients from the general population, predictors of mortality were being older than 60 years (the risk increasing

Table 1 Univariable analyses to predict mortality (derivation samples)

Variables	Nursing home residents (N=2140)		General population (N=9121)	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
Gender				
Female	Ref	–	Ref	–
Male	1.50 (1.22–1.85)	0.0002	2.81 (2.31–3.42)	<0.0001
Age (years) categorized				
≤59	Ref	–	Ref	–
60–69	Ref	–	10.89 (6.76–17.54)	<0.0001
70–79	1.67 (0.99–2.81)	0.0568	35.48 (22.75–55.32)	<0.0001
80–89	2.40 (1.51–3.81)	0.0002	103.37 (66.86–159.83)	<0.0001
≥90	3.23 (2.04–5.13)	<0.0001	142.31 (86.55–234)	<0.0001
Hospital admission	4.01 (3.23–4.99)	<0.0001	24.13 (17.51–33.25)	<0.0001
Previous hospital admissions 1 month	1.93 (0.97–3.86)	0.0631	13.16 (9.73–17.79)	<0.0001
Comorbidities				
Cardiovascular	1.66 (1.36–2.04)	<0.0001	10.10 (8.30–12.30)	<0.0001
Cerebrovascular	1.36 (1.11–1.67)	0.0032	5.64 (4.53–7.04)	<0.0001
Hemiplegia / Paraplegia	0.65 (0.43–0.97)	0.0349	4.78 (2.83–8.08)	<0.0001
Dementia	1.97 (1.60–2.41)	<0.0001	13.52 (9.61–19.02)	<0.0001
Respiratory (including COPD)	1.23 (0.97–1.56)	0.0831	3.47 (2.86–4.22)	<0.0001
Liver disease				
No	Ref	–	Ref	–
Mild	1.30 (0.90–1.89)	0.1652	3.36 (2.52–4.49)	<0.0001
Moderate/severe	0.31 (0.07–1.33)	0.1145	9.17 (4.97–16.93)	<0.0001
Diabetes				
No	Ref	–	Ref	–
Yes, without organ damage	1.02 (0.80–1.30)	0.8695	4.76 (3.78–6)	<0.0001
Yes, with organ damage	1.28 (0.85–1.92)	0.2393	13.75 (9.94–19.01)	<0.0001
Kidney	1.71 (1.36–2.15)	<0.0001	9.44 (7.46–11.95)	<0.0001
Cancer	1.11 (0.80–1.55)	0.5254	3.58 (2.90–4.43)	<0.0001
Rheumatic	0.92 (0.59–1.43)	0.7138	3.53 (2.46–5.07)	<0.0001
Peptic ulcer	1.29 (0.88–1.88)	0.1950	4.02 (2.82–5.72)	<0.0001
Inflammatory bowel disease	0.95 (0.35–2.58)	0.9160	0.72 (0.37–1.42)	0.3462
Arterial hypertension	1.30 (1.04–1.62)	0.0221	8.17 (6.61–10.10)	<0.0001
Dyslipidemia	1.11 (0.91–1.35)	0.3236	3.36 (2.78–4.05)	<0.0001
Asthma	0.91 (0.60–1.38)	0.6558	0.94 (0.68–1.29)	0.6843
Basal treatment				
Antidiabetics	0.77 (0.59–1.01)	0.0565	3.35 (2.62–4.28)	<0.0001
Antihypertensive	1.55 (0.72–3.30)	0.2617	5.55 (3.26–9.44)	<0.0001
Diuretics	1.24 (1–1.53)	0.0478	9.53 (7.64–11.88)	<0.0001
Cardiovascular	1.31 (1.01–1.70)	0.0415	5.52 (4.35–7.01)	<0.0001
Beta-blockers	0.87 (0.66–1.14)	0.3126	5.72 (4.52–7.23)	<0.0001
Calcium channel blockers	1.04 (0.83–1.32)	0.7248	7.38 (5.98–9.12)	<0.0001
Renin–angiotensin aldosterone system inhibitors	1.09 (0.89–1.34)	0.3934	7.04 (5.78–8.58)	<0.0001
NSAIDs	0.52 (0.30–0.93)	0.0274	0.32 (0.22–0.46)	<0.0001
Lipid lowering drugs/statins	0.58 (0.45–0.76)	<0.0001	3.59 (2.94–4.37)	<0.0001
Antiplatelets	0.95 (0.76–1.19)	0.6666	6.08 (4.87–7.59)	<0.0001
Direct oral anticoagulants	0.37 (0.26–0.54)	<0.0001	2.71 (1.88–3.91)	<0.0001
Heparin	1.49 (0.72–3.06)	0.2804	4.75 (2.07–10.94)	0.0002
Bronchodilators	1.05 (0.79–1.39)	0.7401	2.41 (1.92–3.03)	<0.0001
Chronic azithromycin	2.72 (1.17–6.33)	0.0204	6.23 (3.23–12.03)	<0.0001

Table 1 (continued)

Variables	Nursing home residents (<i>N</i> =2140)		General population (<i>N</i> =9121)	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
Immunosuppressants	0.25 (0.03–1.89)	0.1778	2.10 (1.15–3.84)	0.0162
Chronic systemic steroids	0.92 (0.54–1.56)	0.7513	3.78 (2.61–5.46)	<0.0001

OR odds ratio, CI confidence interval

for older ages), being male, having been admitted to a hospital in the month before admission for any cause, and being admitted to a hospital for COVID-19, as well as any of the following comorbidities: cardiovascular disease, dementia, respiratory, liver disease, diabetes with organ damage, or cancer. Of the baseline medications considered, use of anti-coagulants was again protective. The AUC for this model was 0.941 (Table 2).

Some of the variables that in the univariable analysis showed a predictive or protective relationship with mortality did not appear in the multivariable models or even reverse their relationship due to the confounding effect with the rest of the variables included in the final models. Additionally, patients with dementia and nursing home were admitted to the ICU to a lesser extent than the general population because of resource limitation. We performed a multivariate logistic regression to study the differences in ICU hospitalization according to whether the patients is nursing home resident or from the general population, and adjusting for dementia, sex and age, and we concluded that the risk of ICU hospitalization is significantly lower in nursing home residents, both among patients with dementia (OR (95% CI)=0.10 (0.01–0.94), $p=0.0443$) and without dementia (OR (95% CI)=0.25 (0.15–0.44), $p<0.0001$)).

For each category of a significant predictor, we assigned a weight and created risk scores for each sample. For the continuous risk scores, the AUCs in the nursing home sample were 0.754 in the derivation and 0.717 in the validation sample and in the general population sample were 0.941 in the derivation and 0.938 in the validation sample. Similar AUCs were obtained by *K*-fold cross validation (0.753 and 0.933, respectively) (Table 2).

Table 3 shows the division of these continuous risk scores into categories indicating a low to a very high risk of dying in the short term. The AUC for the risk categories for nursing home residents was 0.754 in the derivation and 0.714 in the validation sample, while for the general population, it was 0.916 in the derivation and 0.914 in the validation sample.

Figure 1 shows the Kaplan–Meier curves for risk categories in each sample. The log-rank test detected significant differences between all risk categories in both derivation and validation samples, except between those with scores

of 4–5 and 6–10 in the nursing home resident validation sample ($p=0.612$).

Discussion

This study analyses outcomes in a sample of people with laboratory-confirmed SARS-CoV-2 infection from an entire region, separating nursing home residents from the general population, seeking to identify predictors of short-term mortality in a large cohort including patients who were and were not hospitalized. We have succeeded in identifying a set of comorbidities and baseline treatments related to death with a good predictive capacity for people from the general population.

To date, several meta-analyses have explored the relationship between COVID-19 and mortality [19, 20]. In all cases, a potential weakness is heterogeneity in the data, and all these analyses have focused on hospitalized patients, using laboratory test results, which were not uniformly selected and evaluated. Our study confirms previously published findings in that advanced age, male sex and comorbidities are associated with a higher risk of mortality. Additionally, the present study also identifies previous hospitalizations and some chronic baseline treatments as associated with death from COVID-19.

Most of the patients in the general population and nursing homes were elderly men with multiple comorbidities, in agreement with previous studies [2–6, 8]. It has been speculated that older patients with chronic diseases are more likely to die of COVID-19, because age-related alterations in immune function weaken the response to SARS-CoV-2 and hence worsen outcomes [21].

A higher proportion of men than women died and this could be partially explained by the stronger effect of older age among men. Circulating sex hormones in males and females might influence susceptibility to COVID-19 infection, as shown in a previous study, because they modulate the responses of adaptive and innate immunity [22].

Some recent meta-analyses have assessed the prevalence of comorbidities in patients with COVID-19; [2, 8, 19]; however, not all comorbidities have the same strength as a predictive risk factor for mortality. Our study showed that

Table 2 Multivariable analysis for prediction of death using multilevel analysis (derivation sample)

Variables	Nursing home residents (<i>N</i> =2140)				General population (<i>N</i> =9121)			
	OR (95% CI)	<i>p</i>	β coefficient	Weight	OR (95% CI)	<i>p</i>	β coefficient	Weight
Gender								
Female	Ref	–	–	0	Ref	–	–	0
Male	1.62 (1.27–2.07)	0.0001	0.4838	3	1.65 (1.30–2.11)	<0.0001	0.5019	1.5
Age (years) categorized								
≤ 59	Ref	–	–	0	Ref	–	–	0
60–69	Ref	–	–	0	4.99 (3.04–8.20)	<0.0001	1.6081	4.5
70–79	Ref	–	–	0	8.90 (5.60–14.45)	<0.0001	2.1970	6
80–89	1.83 (1.33–2.52)	0.0002	0.6049	4	27.17 (16.93–43.60)	<0.0001	3.3021	9
≥ 90	2.59 (1.86–3.60)	<0.0001	0.9511	7	61.16 (35.27–106.05)	<0.0001	4.1135	11.5
Hospital admission	3.84 (3.04–4.86)	<0.0001	1.3465	9	5.96 (4.17–8.52)	<0.0001	1.7848	5
Previous hospital admissions 1 month					3.13 (2.14–4.57)	<0.0001	1.1404	3
Comorbidities								
Cardiovascular	1.49 (1.18–1.88)	0.0007	0.3996	3	1.54 (1.20–1.98)	0.0008	0.4308	1
Dementia	1.80 (1.44–2.25)	<0.0001	0.5861	4	2.74 (1.84–4.10)	<0.0001	1.0093	3
Respiratory (COPD)					1.49 (1.17–1.91)	0.0013	0.4008	1
Liver disease								
No					Ref	–	–	0
Mild					1.43 (1.01–2.03)	0.0448	0.3578	1
Moderate/Severe					2.47 (1.16–5.26)	0.0192	0.9038	2.5
Diabetes								
No + Yes, without organ damage					Ref	–	–	0
Yes, with organ damage					1.81 (1.23–2.69)	0.0029	0.5956	1.5
Kidney	1.33 (1.03–1.72)	0.0283	0.2867	2				
Cancer					1.75 (1.34–2.27)	<0.0001	0.5569	1.5
Basal treatment								
Lipid lowering drugs/statins	0.63 (0.47–0.84)	0.0019	-0.4627	-3				
Direct oral anticoagulants	0.34 (0.23–0.51)	<0.0001	-1.0702	-7	0.34 (0.22–0.52)	<0.0001	-1.0692	-3
AUC (95% CI)*	0.754 (0.730–0.778)				0.941 (0.934–0.948)			
Risk score, range					-10 to 28			
AUC (95% CI)								
Derivation sample	0.754 (0.730–0.778)				0.941 (0.934–0.948)			
K-fold cross validation	0.753 (0.729–0.777)				0.933 (0.923–0.943)			
Validation sample	0.717 (0.687–0.748)				0.938 (0.928–0.947)			

Due to the high mean age of the nursing home group and that there were less than 10% of patients under 70 years, we took as reference group for that sample those < 70 years as a whole

OR odds ratio, CI confidence interval, AUC area under the ROC curve

*The AUC (95% CI) without considering the IHO (integrated healthcare organizations) was 0.742 (0.718–0.766) for the model of patients living in a nursing home, and 0.941 (0.934–0.948) for the model of patients not living in a nursing home

people with underlying cardiovascular disease or dementia are the two groups most likely to die.

The mechanisms underlying the association between cardiovascular disease and COVID-19 might be connected to

infection-related demand ischemia that evolves into myocardial injury and dysfunction and there is evidence of direct viral infection of the myocardium [23]. Regarding dementia as one of the most powerful risk factors for death, our

Table 3 Risk groups of short-term mortality in the derivation and validation samples of both groups using multilevel analysis

Variables	Derivation sample				Validation sample			
	<i>N</i>	Mortality, <i>n</i> (%)	OR (95% CI)	<i>p</i>	<i>N</i>	Mortality, <i>n</i> (%)	OR (95% CI)	<i>p</i>
<i>Nursing home residents</i>								
Score ≤3	475	29 (6.11)	Ref	–	298	20 (6.71)	Ref	–
4 ≤ Score ≤5	262	30 (11.45)	1.98 (1.16–3.38)	0.0127	173	30 (17.34)	2.92 (1.60–5.34)	0.0005
6 ≤ Score ≤10	641	130 (20.28)	3.89 (2.54–5.94)	<0.0001	440	86 (19.55)	3.38 (2.02–5.64)	<0.0001
11 ≤ Score ≤16	511	169 (33.07)	7.61 (5–11.59)	<0.0001	343	119 (34.69)	7.39 (4.45–12.26)	<0.0001
Score ≥17	251	149 (59.36)	23.98 (14.63–36.41)	<0.0001	173	87 (50.29)	14.30 (8.30–24.66)	<0.0001
AUC (95% CI)			0.754 (0.731–0.778)				0.714 (0.684–0.744)	
<i>General population</i>								
Score <10	7152	30 (0.42)	Ref	–	4774	20 (0.42)	Ref	–
10 ≤ Score ≤12.5	861	72 (8.36)	21.49 (13.95–33.11)	<0.0001	574	52 (9.06)	23.50 (13.92–39.68)	<0.0001
13 ≤ Score ≤15	513	102 (19.88)	58.73 (38.62–89.30)	<0.0001	338	70 (20.71)	60.39 (36.16–100.85)	<0.0001
Score >15	593	257 (43.34)	180.21 (121.55–267.17)	<0.0001	391	165 (42.20)	172.77 (106.57–280.09)	<0.0001
AUC (95% CI)			0.916 (0.903–0.929)				0.914 (0.898–0.931)	

OR odds ratio, CI confidence interval, AUC area under the ROC curve

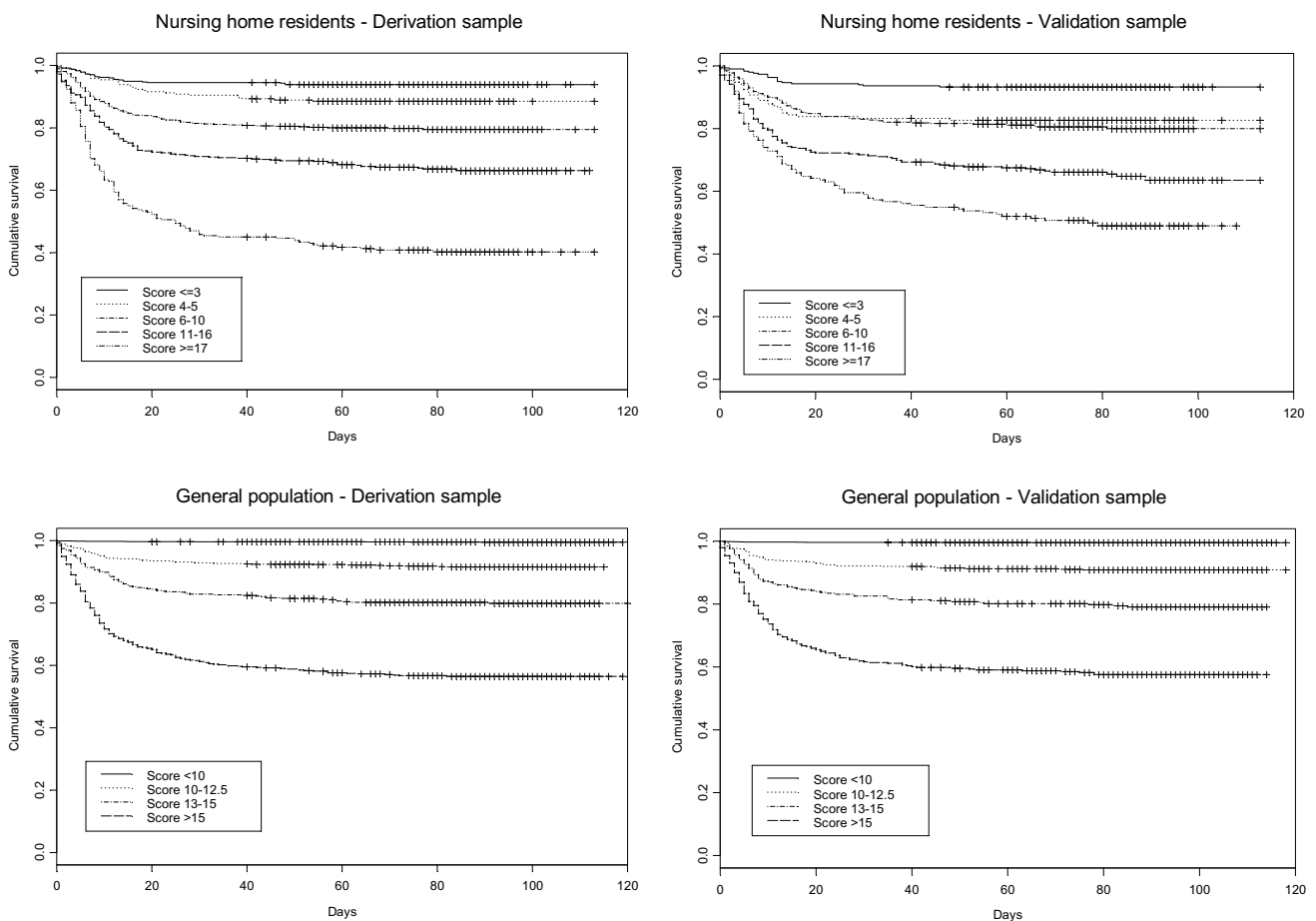


Fig. 1 Kaplan–Meier curves of short-term mortality according to the risk classes in the derivation and validation samples. The log-rank test detected statistically significant differences between all risk

classes in both derivation and validation samples, except between the risk classes with score 4–5 and score 6–10 in the validation sample of nursing home residents ($p=0.612$)

findings are consistent with other studies [24], and it is plausible that respiratory failure, frequent in COVID-19, masks the atypical symptoms in patients with dementia, leading to a failure to recognize the need for medical attention. Furthermore, in this sense patients with dementia and nursing home were admitted to the ICU to a lesser extent than general population because of resource limitation. On the other hand, the physical and cognitive impairment suffered by these patients with loneliness and lockdown worsens their prognosis, so the help of a geriatrician could be valuable.

Comorbidities such as COPD, diabetes, chronic liver disease and cancer were only significant in multivariate analyses for the general population. COPD, inflammation of the lung parenchyma and expiratory airflow limitation may cause respiratory failure, favoring virus superinfection with SARS-CoV-2 [25]. Diabetes is one of the most prevalent underlying conditions in COVID-19 patients. Although the mechanism is not entirely clear, it is suspected that the exacerbated proinflammatory cascade and impaired immune response are involved in this association [26–28]. Despite the low prevalence of chronic liver disease in our patients, consistent with the findings of other studies, this was also associated with higher mortality [29–31]. It seems that patients with this chronic disease are not at greater risk of acquiring the infection, but do have a poorer prognosis once infected.

Patients with cancer are more susceptible to infection because of their systemic immunosuppressive state caused by malignancy and also have a higher risk of mortality [32–34]. Patients with chronic kidney disease are also more vulnerable to COVID-19, and the already impaired kidney function may deteriorate [35, 36].

Hospital admission was associated with a poorer prognosis and higher mortality, as was admission in the previous month. This is clearly related to a more serious presentation of the infection, as reported for other infectious diseases [37].

Another important finding of our study was the protective role of some long-term medications, namely, anticoagulants and statins, as noted by other authors [38–40]. COVID-19 is an inflammatory and prothrombotic disease, and hence, chronic anticoagulation may well provide a real defense against thrombosis [40]. The potential beneficial effects of statins in COVID-19 could be due to their well-known anti-inflammatory properties and might regulate virus replication, exerting a protective effect [38, 39]. The use of statins and anticoagulants increase as age increases up to 89 years; from the age of 90 percentage decrease in both populations (general and nursing home). This decrease could be related to functional and cognitive deterioration of the elderly patients.

Routine prediction rules used in general wards and ICUs are not able to accurately assess the severity and/or mortality

of COVID-19. New validated clinical prediction rules are required for patient stratification [9]. Our rules, based only on variables which are easily accessible and interpretable at the time of diagnosis, can identify seriously ill patients with COVID-19 who are at risk of death. Using data routinely collected in the medical record, we can distinguish patients at high risk (score > 11 for nursing home residents, or > 9 for the general population) from those at low risk. Patients at high risk should be hospitalized and closely monitored, while low-risk individuals could be treated as outpatients under surveillance. To our knowledge, this is the first such prediction rule that achieves this goal. It could help physicians to identify "high-risk groups". These groups should be prioritized if a vaccine becomes available, given the high mortality associated with COVID-19 in combination with these chronic conditions.

Strengths of this study include the large sample size, even for nursing home residents, homogeneity of the data, lack of reliance on data abstractors, avoiding potential bias, and development of predictive models following TRIPOD guidelines [41]. As for limitations of the study, we recognize that the analysis was restricted to a limited number of variables for which we were confident of the validity of the data and we have confirmed this validity. Furthermore, though all cases were COVID-19 positive, it was not verified that the cause of death was unequivocally the SARS-CoV-2 infection in all cases. Finally, this analysis focuses on one region, not the whole of our country or a larger geographical area, and therefore, other studies should be conducted to check the external validity of our models, and thereby, their generalizability.

In conclusion, this study provides for the first time two separate clinical prediction rules for COVID-19 positive individuals from the general population and from nursing homes, using factors related to mortality, that have fairly good predictive value and could be used by general practitioners as they require only basic patient information.

Acknowledgements We are grateful for the support of the Basque Health Service-Osakidetza and the Department of Health of the Basque Government. We also gratefully acknowledge the patients who participated in the study. The authors also would like to thank Ideas Need Communicating Language Services for improving the use of English in the manuscript. The COVID19-Osakidetza Working group includes the following researchers: Susana García-Gutiérrez, Miren Orive, Nerea Gonzalez, Iratxe Lafuente, Ane Anton, Ane Villanueva, Josune Martin, Cristina Muñoz, Maria Jose Legarreta, Urko Aguirre and Jose María Quintana (Research Unit) Hospital de Galdakao-Usansolo; Pedro Pablo España, Ane Uranga, Mikel Egurrola, Amaia Aramburu, Amaia Artaraz, Leire Chasco, Olaia Bronte, Patricia García, Ana Jodar, Virginia Fernandez, and Cristobal Esteban (Respiratory Service) Hospital de Galdakao-Usansolo, Naia Mas(ICU) Hospital de Galdakao-Usansolo, Esther Pulido (Emergency Service) Hospital de Galdakao-Usansolo, Itxaso Bengoetxea (Hospital at Home Service), Galdakao-Barrualde Health Organization; Antonio Escobar, Amaia Bilbao and Iñigo Gorostiza, (Research Unit) Hospital Universitario de Basurto; Iñaki

Arriaga, (Respiratory Service) Hospital Universitario de Basurto; José Joaquín Portu (Internal Medicine) and Naiara Parraza (Research Unit), Instituto Bioaraba; Milagros Iriberrí, Rafael Zalacain, Luis Alberto Ruiz and Leyre Serrano (Respiratory Service) Hospital Universitario de Cruces; Adriana Couto and Oier Ateka (Internal Medicine Service) Hospital Universitario Donostia; Arantza Cano, Hospital Santa Marina (Respiratory Service); Maria Olatz Ibarra, Hospital de Urduliz (Pharmacy Service); Eduardo Millan, Mayte Bacigalupe, Jon Letona and Andoni Arcelay, Central Services Osakidetza; Iñaki Berraondo, Department of Health, Basque Government.

Author contributions JMQ, PPE and SG-G had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: PPE, SG-G, AU, IL, AA-L, AV, MJL, UA, and JMQ. Acquisition, analysis, or interpretation of data: PPE, SG-G, AB, IL, AA-L, AV, AU, MJL, UA, JMQ and the COVID-19-Osakidetza Working group. Drafting of the manuscript: PPE, AB, JMQ. Critical revision of the manuscript for important intellectual content: PPE, AB, AU, SG-G, IL, AA-L, AV, MJL, UA, Quintana and the COVID-19-Osakidetza Working group. Statistical analysis: PPE, AB, AA-L, AV, MJL, UA, JMQ. Obtained funding: SG-G, PPE, JMQ. Administrative, technical, or material support: Muñoz, PPE, JMQ and the COVID-19-Osakidetza Working group.

Funding This work was supported in part by grants from the Instituto de Salud Carlos III and the European Regional Development Fund COVID20/00459; the health outcomes group from Galdakao-Barrualde Health Organization; the Kronikune Institute for Health Service Research; and the thematic network–REDISSEC (Red de Investigación en Servicios de Salud en Enfermedades Crónicas)–of the Instituto de Salud Carlos III.

Availability of data and material Data available on request from the authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study protocol was approved by the Ethics Committee of the Basque Country (reference PI2020059).

Human and animal rights The study protocol was approved by the Ethics Committee of the Basque Country.

Informed consent For this study, at the beginning of the COVID pandemic, formal written consent was not required by the Ethics Committee of the Basque Country due to the special circumstances at that moment.

References

- 2020 WD-GsORatMBoC--M. <https://www.who.int/dg/speeches/detail/who-director-general-sopening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>. Accessed 20 Apr 2020
- Emami A, Javanmardi F, Pirbonyeh N, Akbari A (2020) Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-analysis. *Arch Acad Emerg Med* 8(1):e35
- Guan WJ, Liang WH, Zhao Y et al (2020) Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J*. 55(5):2000547
- Liu H, Chen S, Liu M, Nie H, Lu H (2020) Comorbid chronic diseases are strongly correlated with disease severity among COVID-19 patients: a systematic review and meta-analysis. *Aging Dis* 11(3):668–678
- Wang X, Fang X, Cai Z et al (2020) Comorbid chronic diseases and acute organ injuries are strongly correlated with disease severity and mortality among COVID-19 patients: a systemic review and meta-analysis. *Research (Wash DC)* 2020:2402961
- Yang J, Zheng Y, Gou X et al (2020) Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis* 94:91–95
- Wu Z, McGoogan JM (2020) Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. <https://doi.org/10.1001/jama.2020.2648>
- Zhou F, Yu T, Du R et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395(10229):1054–1062
- Wynants L, Van Calster B, Collins GS et al (2020) Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 369:m1328
- Charlson ME, Sax FL, MacKenzie CR, Fields SD, Braham RL, Douglas RG Jr (1986) Assessing illness severity: does clinical judgment work? *J Chronic Dis* 39(6):439–452
- World Health O, Collaborating Centre for Drug Statistics M (2013) ATC classification index with DDDs. WHO, Oslo
- McFadyen JD, Schaff M, Peter K (2018) Current and future antiplatelet therapies: emphasis on preserving haemostasis. *Nat Rev Cardiol* 15(3):181–191
- Chaudhary R, Sharma T, Garg J et al (2020) Direct oral anticoagulants: a review on the current role and scope of reversal agents. *J Thromb Thrombolysis* 49(2):271–286
- Quan H, Sundararajan V, Halfon P et al (2005) Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 43(11):1130–1139
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1):29–36
- Hastie T, Tibshirani R, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction. Springer, Berlin
- Steyerberg EW (2009) Clinical prediction models a practical approach to development, validation, and updating. Springer, New York
- Barrio I, Arostegui I, Rodríguez-Álvarez MX, Quintana JM (2017) A new approach to categorising continuous variables in prediction models: proposal and validation. *Stat Methods Med Res* 26(6):2586–2602
- Parohan M, Yaghoubi S, Seraji A, Javanbakht MH, Sarraf P, Djalali M (2020) Risk factors for mortality in patients with Coronavirus disease 2019 (COVID-19) infection: a systematic review and meta-analysis of observational studies. *Aging Male* 2020:1–9
- Tian W, Jiang W, Yao J et al (2020) Predictors of mortality in hospitalized COVID-19 patients: a systematic review and meta-analysis. *J Med Virol*. <https://doi.org/10.1002/jmv.26050>
- Opal SM, Girard TD, Ely EW (2005) The immunopathogenesis of sepsis in elderly patients. *Clin Infect Dis* 41(Suppl 7):S504–512
- Jaillon S, Berthenet K, Garlanda C (2019) Sexual dimorphism in innate immunity. *Clin Rev Allergy Immunol* 56(3):308–321
- Guo T, Fan Y, Chen M et al (2020) Cardiovascular implications of fatal outcomes of patients with coronavirus disease 2019 COVID-19. *JAMA Cardiol*. <https://doi.org/10.1001/jamacardio.2020.1017>

24. Bianchetti A, Rozzini R, Guerini F et al (2020) Clinical Presentation of COVID-19 in Dementia Patients. *J Nutr Health Aging* 24(6):560–562
25. Zhao Q, Meng M, Kumar R et al (2020) The impact of COPD and smoking history on the severity of COVID-19: a systemic review and meta-analysis. *J Med Virol*. <https://doi.org/10.1002/jmv.25889>
26. Filardi T, Morano S (2020) COVID-19: is there a link between the course of infection and pharmacological agents in diabetes? *J Endocrinol Invest*. <https://doi.org/10.1007/s40618-020-01318-1>
27. Shi Q, Zhang X, Jiang F et al (2020) Clinical characteristics and risk factors for mortality of COVID-19 patients with diabetes in Wuhan, China: a two-center retrospective study. *Diabetes Care* 43(7):1382–1391
28. Scheen AJ, Marre M, Thivolet C (2020) Prognostic factors in patients with diabetes hospitalized for COVID-19: findings from the CORONADO study and other recent reports. *Diabetes Metab* 46(4):265–271. <https://doi.org/10.1016/j.diabet.2020.05.008>
29. Garrido I, Liberal R, Macedo G (2020) Review article: COVID-19 and liver disease-what we know on 1st May 2020. *Aliment Pharmacol Ther* 52(2):267–275
30. Mantovani A, Beatrice G, Dalbeni A (2020) Coronavirus disease 2019 and prevalence of chronic liver disease: a meta-analysis. *Liver Int* 40(6):1316–1320
31. Jothamani D, Venugopal R, Abedin MF, Kaliamoorthy I, Rela M (2020) COVID-19 and the liver. *J Hepatol* 73(5):1231–1240. <https://doi.org/10.1016/j.jhep.2020.06.006>
32. Extnance A (2020) Covid-19 and long term conditions: what if you have cancer, diabetes, or chronic kidney disease? *BMJ* 368:m1174
33. Liang W, Guan W, Chen R et al (2020) Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol* 21(3):335–337
34. Meng Y, Lu W, Guo E et al (2020) Cancer history is an independent risk factor for mortality in hospitalized COVID-19 patients: a propensity score-matched analysis. *J Hematol Oncol* 13(1):75
35. Li J, Li SX, Zhao LF, Kong DL, Guo ZY (2020) Management recommendations for patients with chronic kidney disease during the novel coronavirus disease 2019 (COVID-19) epidemic. *Chronic Dis Transl Med* 6(2):119–123
36. Cheng Y, Luo R, Wang K et al (2020) Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int* 97(5):829–838
37. Esteban C, Quintana JM, Moraza J et al (2009) Impact of hospitalisations for exacerbations of COPD on health-related quality of life. *Respir Med* 103(8):1201–1208
38. Reiner Ž, Hatamipour M, Banach M et al (2020) Statins and the COVID-19 main protease: in silico evidence on direct interaction. *Arch Med Sci* 16(3):490–496
39. Rodrigues-Diez RR, Tejera-Muñoz A, Marquez-Exposito L, Rayego-Mateos S, Santos Sanchez L, Marchant V, Tejedor Santamaria L, Ramos AM, Ortiz A, Egido J, Ruiz-Ortega M (2020) Statins: could an old friend help in the fight against COVID-19?. *Br J Pharmacol* 177(21):4873–4886. <https://doi.org/10.1111/bph.15166>
40. Sivaloganathan H, Ladikou EE, Chevassut T (2020) COVID-19 mortality in patients on anticoagulants and antiplatelet agents. *Br J Haematol* 190(4):e192–e195. <https://doi.org/10.1111/bjh.16968>
41. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 350:g7594. <https://doi.org/10.1136/bmj.g7594>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.